Anatomical Mirroring: Real-time User-specific Anatomy in Motion Using a Commodity Depth Camera

Armelle Bauer [2,3], Ali-Hamadi Dicko [2,4], François Faure [2,4], Olivier Palombi [1,2,4], Jocelyne Troccaz [3]
[1] LADAF, [2] LJK, [3] TIMC-IMAG, [4] AnatoScope — **INRIA**, **CNRS**, **Univ. Grenoble Alpes**

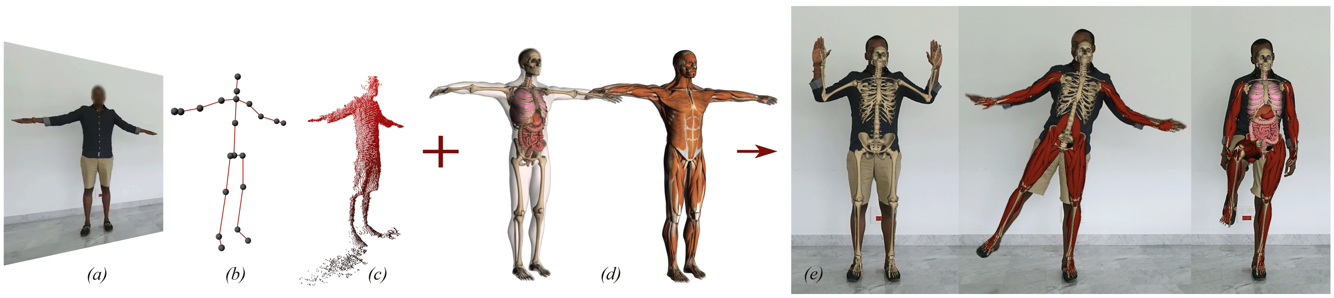

Figure 1: *Using a single Kinect sensor output (color map (a), motion capture skeleton (b), point cloud (c)) and a generic 3D anatomical model (d), a user-specific anatomy is generated and animated in real-time (e).*

Abstract

This paper presents a mirror-like augmented reality (AR) system to display the internal anatomy of a user. Using a single Microsoft V2.0 *Kinect*, we animate in real-time a user-specific internal anatomy according to the user's motion and we superimpose it onto the user's color map, as shown in Fig.*1.e*. The user can visualize his anatomy moving as if he was able to look inside his own body in real-time.

A new calibration procedure to set up and attach a user-specific anatomy to the *Kinect* body tracking skeleton is introduced. At calibration time, the bone lengths are estimated using a set of poses. By using *Kinect* data as input, the practical limitation of skin correspondance in prior work is overcome. The generic 3D anatomical model is attached to the internal anatomy registration skeleton, and warped on the depth image using a novel elastic deformer, subject to a closest-point registration force and anatomical constraints.
The noise in *Kinect* outputs precludes any realistic human display. Therefore, a novel filter to reconstruct plausible motions based on fixed length bones as well as realistic angular degrees of freedom (DOFs) and limits is introduced to enforce anatomical plausibility. Anatomical constraints applied to the *Kinect* body tracking skeleton joints are used to maximize the physical plausibility of the anatomy motion, while minimizing the distance to the raw data. At run-time, a simulation loop is used to attract the bones towards the raw data, and skinning shaders efficiently drag the resulting anatomy to the user's tracked motion.

Our user-specific internal anatomy model is validated by comparing the skeleton with segmented MRI images. A user study is established to evaluate the believability of the animated anatomy.

Keywords: User-specific anatomy, Augmented Human, Real-Time, Motion Capture, Augmented Reality, Markerless Device.

Concepts: •**Computing methodologies** → **Motion capture; Mixed / augmented reality;**

Introduction

The emergence of commodity depth cameras such as *Kinect* sensors motivates new educational, medical and healthcare applications. However, previous studies show that raw *Kinect* data cannot be easily employed in human motion tracking [Pfister et al. 2014; Malinowski and Matsinos 2015]. In this paper, a new calibration and motion capture sufficiently accurate for AR applications are introduced and demonstrated by superimposing internal anatomy on the user's color map in real-time.
At calibration time, the length and width of body segments are estimated based on specific body poses and silhouettes. A novel anatomically sound deformer is applied to fit a high-quality generic 3D biomechanical model in order to generate a user-specific anatomical model. At run-time, our model tracks bone motions based on the *Kinect* body joints output, while enforcing anatomical plausibility rules such as constant lengths and joint limits. Our user study preliminary shows that the precision is sufficient to superimpose the user-specific anatomical model onto the color image, using linear blend skinning.

The paper is organized as follows: *Section 1* briefly survey related work. *Section 2* introduces a body size measurement procedure based on multiple poses and silhouette points and an anatomically sound deformer well adapted to *Kinect* outputs. *Section 3* describes how it is animated based on robust motion capture using anatomical constraints. *Section 4* goes through results and the validation process. *Section 5* finally concludes by presenting possible applications of this work and under development features.

1 Related Work

Nowadays, human body modeling and tracking are widely studied for a variety of applications such as motion capture or morphometric studies.

Skin Registration is the most accessible approach to generate a wide range of human bodies. Most studies are based on skin statistical models generated by a shape and pose database [Helten et al. 2013]. [Gilles et al. 2011] use frame-based skinning methods to deform a generic skin to fit at best the user data. Other approaches using point cloud [Li et al. 2013] or multiposition silhouettes [Vlasic et al. 2008] may also be used to reconstruct the body skin. Most often, raw data come from acquisition of people wearing clothes and this may lead to non-realistic bodies (part proportions, *etc*). [Bălan and Black 2008] and [Zeng et al.

2015] intend to find ways to pass through these limitations. Since they rely only on skin models and do not include internal anatomy, those methods may result in unrealistic skin twisting.

Anatomy Registration The most accurate subject-specific anatomy registration methods come from the medical imaging fields [Sotiras et al. 2012]. However, 3D medical images are not easily used in a non medical context and are not adapted to real-time capture. Several other methods have been proposed. [Quah et al. 2005] present a pose-dependent method to register a 3D anatomical model onto 2D images. Based on key points, they register skin and skeleton (no soft tissues). However this method gives static results. Using *Kinect* point cloud, [Zhu et al. 2015] register user-specific skin and skeleton during motion.

[Dicko et al. 2013] as [Gilles et al. 2010], present a pose-dependent method to transfer 3D anatomical model to a target skin. This method is static and time consuming.
The method introduced by [Saito et al. 2015] achieves near-interactive run time skin, skeleton and soft tissue 3D model editing.

Fig.*2* summerizes this comparison between state of the art anatomy registration methods.

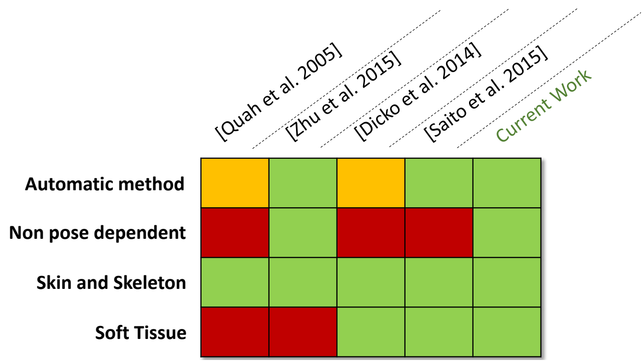

Figure 2: *Comparison between state of the art anatomy registration methods. Legend : green means that the characteristic is totally handled by the method, red that it is not, and orange that it is partly handled.*

User Tracking In [Pfister et al. 2014] the authors assess that the basic *Kinect* body tracking results are enough for basic motion measurements such as stride timing, joint angles during motion, but are far beyond Vicon cameras in terms of software and hardware.

The tracking algorithm used in this paper is based on the *Kinect* SDK animated skeleton which is really noisy. Whereas we add constraints to upgrade the tracking, [Meng et al. 2013] ask the user to pinpoint anatomical key points to help positioning the data. [Shen et al. 2012] use an example-based method to learn how to correct initially tracked poses.

Because body tracking is a critical step, other methods like [Zhou et al. 2014] or [Wei et al. 2012] use the *Kinect* depth map and implement their own posture registration process using probabilities or pose estimations. [Zhu et al. 2015] use multi-*Kinect* depth maps and anatomical knowledge to enhance realistic limb motions.

Nowadays, in the game industry, sports and fitness training applications using depth map tracking devices are commonly used (eg. Nike *Kinect*+, Get Fit With Mel B, Your Shape, *etc...*). To our knowledge, the best tracking games are based on the *Kinect* technology. All of these games shows the user depth map or silhouette only.

With AR, a precision and a realism constraint are added compared to this field state of the art by presenting the anatomy superimposed onto the user's color map. Fig.*3* highlights *Kinect* tracking problems such as disconnected bones head and overlaps between bones.

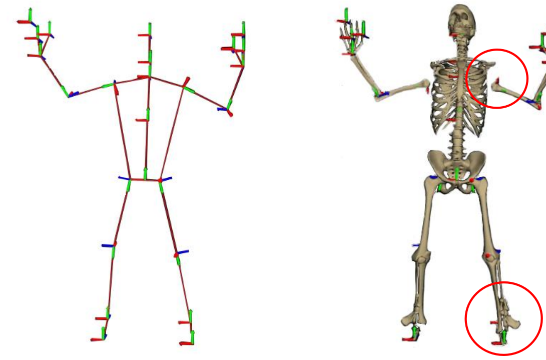

Figure 3: *Kinect body tracking system and user-specific anatomy.*

AR Systems In the last few years, the number of AR applications increased in the medical education field [Kamphuis et al. 2014].

The Magic Mirror [Blum et al. 2012] superimposes statically CT scans of the abdomen onto the user's image. The Digital Mirror [Maitre 2014] shows full body CT scans but does not superimpose them on the user image. In these two cases, data follow the user's motion but are not deformed with respect to these motions.

The Anatomical Mirror [Borner and Kirsch 2015] allows full-body motion by using the *Kinect* tracking, but it displays animated generic 3D models while we show a user-specific one.

Thanks to the use of anatomical knowledge, we significantly improve AR realism and anatomy motion plausibility with respect to [Bauer et al. 2014] and [Bauer et al. 2015] in the Living Book of Anatomy project.

Fig.*4* summerizes comparisons between state-of-the-art demos and our work.

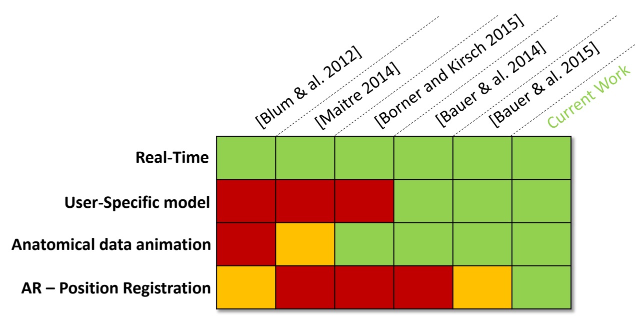

Figure 4: *Comparison between anatomical mirror-like applications. Legend : green for good, orange for average, red for bad.*

Data Validation Validation of anatomical data requires in-vivo measurements, the simplest way is to use as ground truth body measurements [Dao et al. 2014] or/and anatomical landmarks [Espitia-Contreras et al. 2014] taken directly onto the user's body. The study made by [Malinowski and Matsinos 2015] gives limb bones length during motion and compare them with ground truth body measurements.

Using user body anatomical landmarks introduces measurement errors due to body position and skin curvature. We decided to use MRI data as ground truth: in addition to externally visible specific anatomical points to be able to obtain internal specific points (eg. femoral head of bone).

2 User-Specific Anatomy

We present a novel approach using *Kinect* SDK outputs (color map, body tracking skeleton and point cloud) and a 3D reference model including skin surface and internal anatomy (skeleton, muscles, organs, etc) to generate user-specific anatomical data.

The method consists of four steps. First, the user-specific body segment lengths and widths are computed using the *Kinect* SDK outputs (see *Section 2.1*) to define a list of 3D key points. In the second step the generic skin is deformed based on key points and the partial user's point cloud (*Section 2.2*). The third step consists in transfering the reference skeleton inside the user-specific skin (*Section 2.3*). Finally the soft tissue between the bones and the skin is determinded using Laplacian interpolation in a way similar to [Dicko et al. 2013]. These different steps are summarized in Fig.*5*.

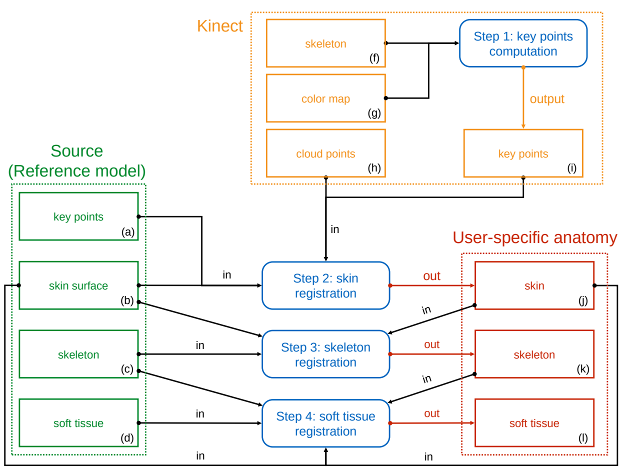

Figure 5: *Pipeline of the user-specific anatomy generation.*

To ease the understanding of the rest of this section, descriptions of each type of deformation skeleton used are provided below:

- ***Kinect* body tracking skeleton**: composed of 25 joints, this animation skeleton is given by the *Kinect* SDK.
- **skin registration skeleton**: composed of 22 control frames (Fig.*8, red dots*) defined on the generic 3d skin and corresponding to some of the *Kinect* body tracking joints, and 18 control points defined on the generic 3d skin contour (Fig.*8, green dots*) that corresponds to the silhouette key points computed in *Section 2.2*. This system is used for skin registration.
- **Internal anatomy registration skeleton**: composed of 96 joint contraints between bones and 373 control frames position (see Fig.*11*). This system is used to keep anatomical consistency during internal skeleton registration and is defined based on anatomical rules presented in *Section 2.3*.

2.1 Body size

The *Kinect* SDK provides a simple body tracking skeleton, without temporal coherence: links in-between segments may have different lengths at each frame. At calibration time: starting from a T-pose (Fig.*6.a*), the user flexes his or her elbows (Fig.*6.b*) and knees (Fig.*6.c*). This allows us to estimate the lengths of upper limb and lower limb segments.

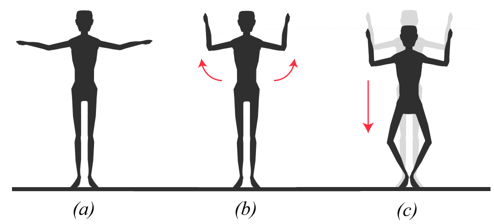

Figure 6: *Calibration process: (a) to have global proportions (head and torso), (b) for real upper limb parts lengths, (c) for real lower limb parts lengths.*

The user silhouette and the body tracking skeleton given by *Kinect* are needed to compute body measurements (see Fig.*7.b*) and define the 18 key points used for skin registration, as presented in *Section 2.2*.
The *Kinect* body tracking skeleton is mapped from camera space to image space using *Kinect* SDK tools.

A key point corresponds to the intersection between the user's silhouette edge pixel and a perpendicular line computed using a Bresenham algorithm. For robustness, we have designed a silhouette detection criterion: an edge pixel is defined by a black pixel followed by three white pixels to avoid silhouette holes.

For each key point, the Bresenham algorithm is initialized using the middle of in-between link segments as starting point and the perpendicular vector as the direction to follow. For instance using the point in-between shoulder and elbow link gives us upper arm width.

The 2d key points found are mapped from image space to camera space using *Kinect* SDK tools, Fig.*7.c* shows the key points we use (eg. body tracking joints points, silhouette head points, silhouette waist points, etc. ...).

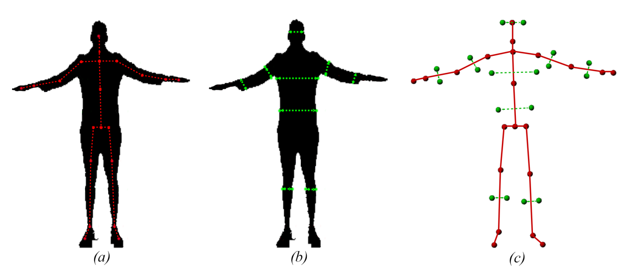

Figure 7: *(a): skeleton key points. (b): body measurements key points. (c): 3D key points used in skin registration.*

Due to clothing and occlusion, some dimensions might be unreliable, especially thigh widths. Firstly, by assuming the human body symmetric along the sagittal plane, small errors in limb lengths are avoided. For each limb the average length value is used as real length in both sides. Other key point positions are infered based on the user silhouette and basic anatomical knowledge. Based on an average human body, we defined ratios between body parts. For instance, knowing that the thigh measurement should be half of the hip measurement, the thigh width can be infered. Some validations are shown in *Section 4*.

2.2 Skin registration

The skin registration method is based on the silhouette key points computed in *Section 2.1* and the *Kinect* point cloud. The main difficulties are the inaccuracy of the *Kinect* output data and the fact that people clothes are captured into the *Kinect* point cloud. To solve these issues, a new elastic deformer is introduced. The skin is rigged using frame-based elastic deformers [Gilles et al. 2011] corresponding to the *Kinect* body tracking skeleton joints (red dots in Fig.*8*). Each skin vertex is controled by several frames, using linear blend skinning. The skinning weights are computed using Voronoi shape functions as in [Faure et al. 2011]. The silhouette key points (green dots in Fig.*8*) are mapped onto the skin to optimize the final result.

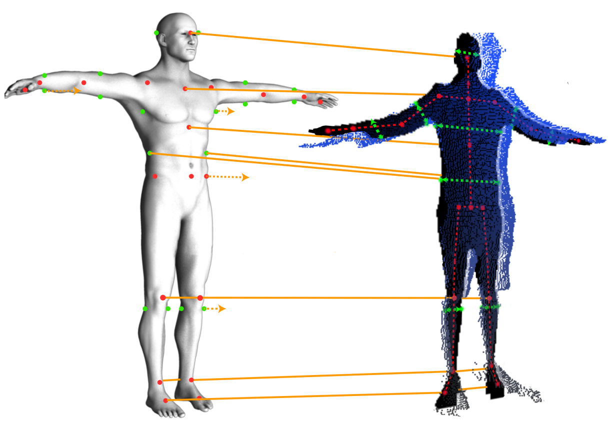

Figure 8: *Skin registration. Red dots: origins of control frames; green dots: silhouette key points; blue dots: Kinect point cloud.*

Instead of using global affine transformations (12DOFs) as in [Dicko et al. 2013]; we use 9DOFs scalable rigids as frames, each bone matrix combines 3 translation, 3 rotation and 3 scale. The advantage over affine control frames is obtaining a better non-uniform local scaling to avoid shearing artefacts.

The skin model is registered to the target by minimizing a weighted sum of three energies [Gilles et al. 2011; Gilles et al. 2013] using an implicit solver.

The predominant energy $E_{skeleton}$ attracts the control frames of the template to the bones of the user-specific model (red points in Fig.*8*). Then the energy $E_{keypoint}$ attracts the silhouette points (green points in Fig.*8*). Minimizing these two first energies scales the limbs, the torso, the neck and the head of the generic model according to the target body measurements as illustrated in Fig.*9.a* and *b*.
The energy $E_{cloudpoint}$ attracts the skin to the target point cloud using an ICP approach to define the correspondance. The forces are propagated from the skin vertices to the degrees of freedom: skeleton control frames (see Fig.*8* and Fig.*9.c*).

Thanks to the fact that a small set of control frames are used, awkward configurations are avoided and no smoothness or kinematic constraint terms are needed.

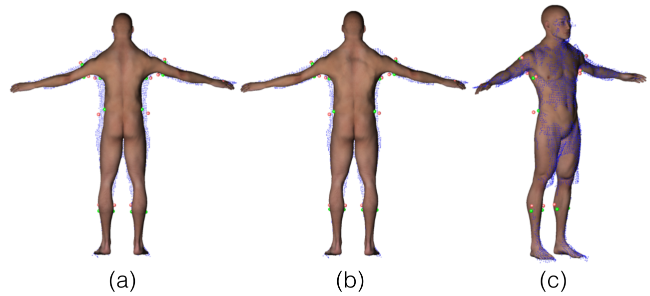

Figure 9: *Skin registration result at the end of each step of the optimization process. (a): minimizing $E_{skeleton}$. (b): minimizing $E_{skeleton}$ and $E_{keypoint}$. (c): minimizing the three energies.*

Fig.*10* presents the skin results after registration with the corresponding *Kinect* point cloud. By using $E_{cloudpoint}$, the torso skin is slightly deformed to refine the model: in the same way, the user being a woman or a man.

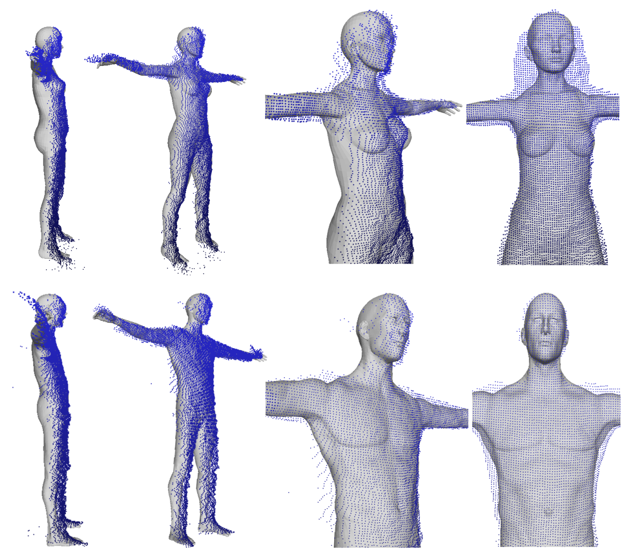

Figure 10: *Kinect point cloud and corresponding registered skin. Top: 1.55m female. Bottom: 1.85m male.*

2.3 Internal Anatomy Registration

User-specific anatomy reconstruction is divided in two sub-parts: anatomical skeleton registration and soft tissue registration. Soft tissue are deformed as described in [Dicko et al. 2013]; here the only focus is on internal skeleton registration. Inputs are the 3D reference of skin and skeleton model and the estimate of the user skin registered obtained in *Section 2.2*.

First, our method uses a volumetric interpolation to estimate the user anatomical skeleton. As in [Dicko et al. 2013], the use of Laplacian interpolation (Fig.*12.a*) with as boundary condition the transformations between the two skins ensures that all the internal anatomy is bounded inside the user's skin after transfer.

A major limitation of the Anatomy Transfer [Dicko et al. 2013] is the fact that the joint structure of the generic model is not maintained. Nothing prevents a bone from passing through another one (Fig.*12.b*) or from being disconnected from a bone to which it should be connected (*for instance ribs and thoracic vertebra, or ulna and humerus around the elbow joint*, see Fig.*12.c*). To keep correct joint structures and avoid these issues, joint constraints between the elements of our elastic bone model are added. The joint location, kinematics and limits are set according to [Nordin and Frankel 2001] (see Fig.*11*).

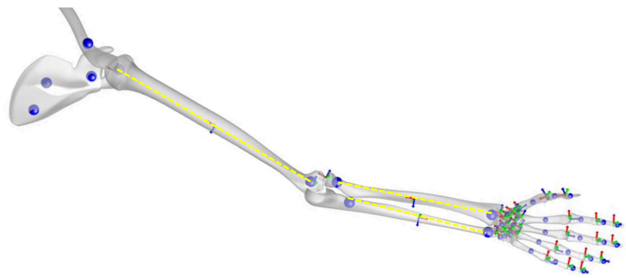

Figure 11: *Right arm internal anatomy registration skeleton. Blue dots for control frame positions; yellow lines and middle bone frames for alignment constraints; other frames for joint contraints.*

Thus, the internal anatomy registration skeleton is defined using frame based elastic deformations [Gilles et al. 2010] with weights computed using a Voronoi shape function as in [Faure et al. 2011] to smoothly propagate all long the bone each control frame transformation. 9DOFs scalable rigids for the control are used to keep head bone consistency as it is in the generic model. This guarantees that the bone heads can only translate, rotate and scale, and thus they keep a similar type of shape as in the generic bone model.
The list of anatomical rules used to define the internal anatomy registration skeleton follows:

- *R01*: Keep long bones straightness (no bending or twisting)
- *R02*: Keep 3D model consistency: the complete set of entities is transferred to avoid holes
- *R03*: Keep bone head consistency
- *R04*: Keep consistency of rib cage and limbs: symmetry with respect to the sagital plane
- *R05*: Keep body joints consistency: type of joint and movement amplitude

To avoid bending bones (Fig.*12.d*), an alignment constraint is added between the two bone heads. This constraint restrains the possible displacements between the control frames in only one direction defined by the line between them (see yellow lines in Fig.*11*). Thereby, the control frames can translate in one direction, but can still scale in all three directions. This alignment constraint is applied to long bones only.

It has been shown in [Zhu et al. 2015] and similar approaches has been explored in [Saito et al. 2015] that non-uniform scaling can be used to get more plausible bone deformations. This is why we introduced more control frames per anatomical bone. The number of frames varies according to bone type, the goal being to give enough deformability to each (*for the registration process*) while keeping good computation times (see blue dots in Fig.*11*). For the short bones such as carpal bones, one frame per bone is used. For the long bones such as the femur two frames per bone are needed: one at the center of each bone head. For the flat bones such as the ribs

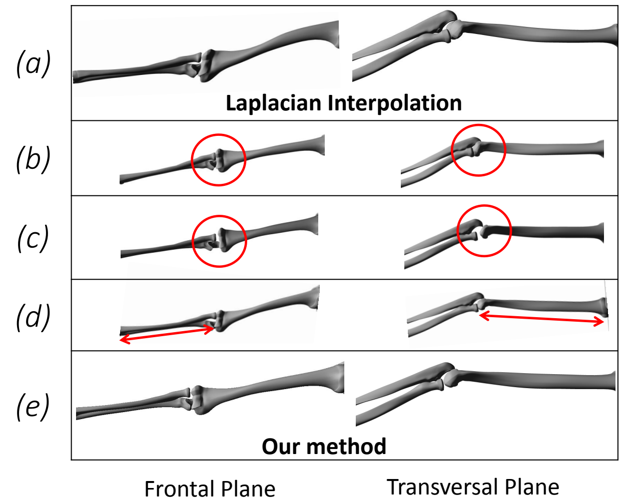

Figure 12: *(a): Laplacian Interpolation, (b): registration without joint contraint (overlaps between bones), (c): registration without joint contraint (bone heads disconnected), (d): registration without alignment constraint (bent and twisted bones), (e): our method.*

three frames per bone are defined to keep ribs close to the skin in terms of curvature: two on bone heads (*eg. close to the joints rib-vertebra and rib-sternum*), and one between the two others (*middle of the rib*). For bones with more complex shape such as vertebrae three frames per bone allows enough deformability to register the model while avoiding overlaps (*e.g overlaps between facet joints, and spinous process of two different vertebrae*). The complexity of the skull deserves a special treatment: use of five control frames for the whole skull deformation.

3 User Tracking

A single *Kinect* (markerless depth sensor) is used to perform body tracking. To reduce tracking noise, we record *Kinect* data in daylight, *Kinect* gives better results with background and ground matte materials. We observed that if the user's ground reflection is too visible, the *Kinect* includes it as part of the user silhouette which leads to lower limb length errors. The *Kinect* position is 60cm off ground for good lower-limb tracking results as determined in [Pfister et al. 2014].

Because *Kinect* segments the depth map to compute body tracking joints at each frame, the in-between link distances change from frame to frame. This leads to disconnected articulations anatomical skeleton (on the limbs) or elongated meshes (on the torso zone).

We present the pipeline of our enhanced body tracking system in Fig.*13*. Firstly we define a hierarchical body tracking system by constraining the limb lengths and by recomputing joint orientations (see *Section 3.1* for more details).
To smooth out small tracking noise, we apply a *Kalman* filter onto the joint positions. Joint orientations are recomputed from the filtered joint positions.
Then we anatomically constrain the joint orientations: more details are given in *Section 3.2*.

3.1 Hierarchical body tracking system

Our hierarchical body tracking system is composed of 25 joints according to the *Kinect* SDK body tracking system.
To define each joint f, the position and the orientation of its par-

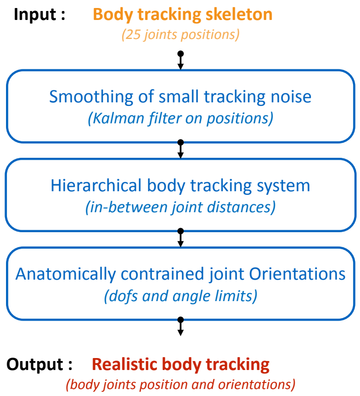

Figure 13: *Enhanced body tracking pipeline.*

ent p is required. To overcome this, we begin by computing the joints from the root (spine base joint) to the leaves (eg. hand tips, foot joints and head joint). The root joint is defined by keeping the filtered *Kinect* position and orientation.

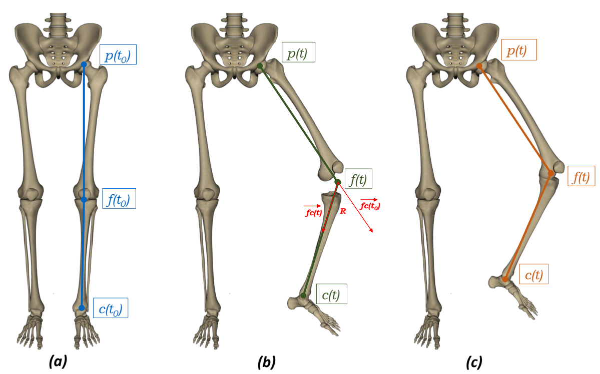

Figure 14: *(a): our hierarchical body tracking skeleton at (t_0). (b): Kinect body tracking skeleton at (t). (c): our result.*

At initialization time t_0 (see Fig.*14.a*), each joint position is defined by our generic anatomical model and in-between link distances computed after calibration (see *Section 2*); and each joint orientation is defined by the initial *Kinect* orientation determined in *Kinect* SDK.

The advantage of using a hierarchical skeleton is to obtain the body pose at each time t using only the joint rotations. We use the current *Kinect* body tracking skeleton to retrieve these rotations.
Most of the time, orientations given by *Kinect* are incorrect so we decided to recompute them using in-between link directions by finding the smallest rotation R between initial direction ($fc(t_0)$) and current direction ($fc(t)$), see Fig.*14.b*. The 3x3 rotation matrix R is defined by the rotation of angle α around $axis$. For more details, see equation 1 and 2.

$$\alpha = asin\left(\left\|\overrightarrow{fc}(t_0) \wedge \overrightarrow{fc}(t)\right\|\right) \tag{1}$$

$$axis = \frac{\left(\overrightarrow{fc}(t_0) \wedge \overrightarrow{fc}(t)\right)}{\left\|\overrightarrow{fc}(t_0) \wedge \overrightarrow{fc}(t)\right\|} \tag{2}$$

Fig.*14.c* shows our hierarchical body skeleton system at step t.

3.2 Anatomically contrained joint orientations

To correct non-anatomically plausible behaviors due to tracking errors, each *Kinect* hierarchical body tracking joint orientation is constrained by limiting the number of possible rotations based on anatomical motion knowledge (eg. knee joint can be aproximated as a 1DOF joint, whereas the hip joint is a 3DOFs joint). This is done by constraining a given quaternion using Euler-angle constraints to find the closest rotation matrix defined only with valid axis within the joint limits. Computation is made using the *Geometric Tools* library [Eberly 2008].
Fig.*15.a* illustrates in red a raw *Kinect* tracking and in grey the result after applying this constraint. To add even more anatomical plausibility to the result, joint limits are added to each rotation axis. Fig.*15.b* highlights this constraint by showing *Kinect* raw data in red and realistic angular limits obtained in grey.

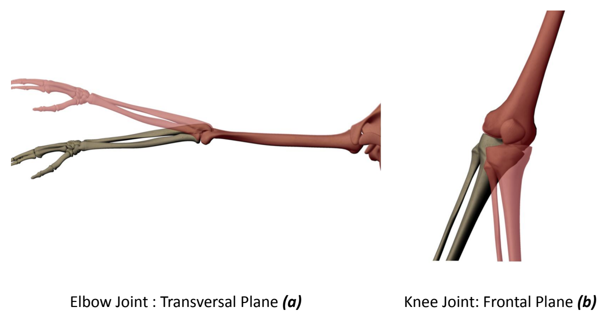

Figure 15: *Kinect data in red and corrected in grey. (a): off angular limits rotation. (b): rotation axis error (DOFs).*

4 Results and Validation

To our knowledge, dealing with realistic anatomy visualization and motion is one of the most complex AR system ever because superimposing 3D anatomical data onto the user's color map reveals all the user measurement and tracking errors.

Our calibration method is a little time consuming (1-2*sec* for skin registration,15-30*sec* for skeleton registration and 30-60*sec* for soft tissue registration) but allows us to obtain a 3D model with accurate user measurements; moreover the motion capture pipeline, even with the introduction of delay during quick motions, leads to realistic and stable user tracking.

Thanks to these two features, the presented method allows a realistic experience for understanding anatomy. The described method is implemented in C++ and runs on a commodity laptop (Intel CoreI7 processor at 3 GHz, Nvidia Quadro K2100M and 8GB of RAM). The real-time AR visualization runs between 35 to 62 fps depending on the 3D feedback: full-body musculoskeletal system (49211 vertices, 95189 faces) will run at 35 fps whereas internal organs (20144 vertices, 39491 faces) will run at 62 fps.
The computational bottleneck of our system is the quality of the 3D model (number of faces and vertices) alongside the quality of the user color map (*Kinect* gives a high definition color map, which is reloaded at each frame).

The visual feedback can be performed on a commodity laptop screen or can be projected onto a 1.50*m*/2.0*m* screen for a demo display (see right side of Fig.*16*).

Fig.*16* presents snapshots of the provided visualization. In a first set of experiments, the motion sequences were acquired for 4 men with an average height of 1.70m, and 3 women with an average height of 1.60m. To get uniform results we work with *Kinect* sequences made

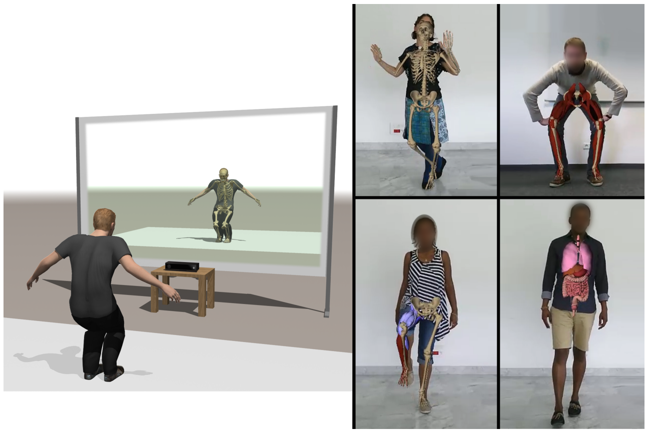

Figure 16: *Left: system set-up. Right: snapshots of results.*

in similar environment conditions (daylight, background material reflections, *Kinect* position, *etc*...).

Fig.*17* presents two tracking data of the same user wearing different clothing and with different hair styles. It can be seen on the right side that the registered skeleton for these two datasets are almost identical; the red one is a little bigger (1.2% in the limbs lengths and 2.5% in torso widths) than the other one (green). This comparison allows the validation of our skin registration process (see *Section 2.2*).

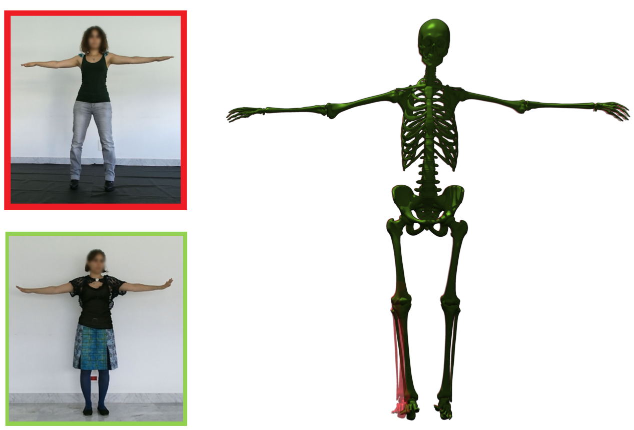

Figure 17: *For the same user with different clothing and hair style (Left), we obtain almost identical results (Right).*

4.1 Validation with MRI

The major contribution of our work, and also the most critical point is the closeness between the user-specific anatomy generated and the user's own. As explained in *Section 1*, using MRI data as ground truth allows us to obtain external as well as internal specific anatomical points for validation purpose.

Fig.*18* present MRI data of two users in front and lateral views side to side with the corresponding 3D user-specific registered anatomies. The internal anatomy registration skeleton introduced in *Section 2.3* is used to set the 3D model in a similar pose as user's in MRI data. After comparing body height, we found an average error of 1.5% between 3D and MRI data which is quite accurate. 3D data being always smaller than MRI data, this error is due to limited skull deformations.

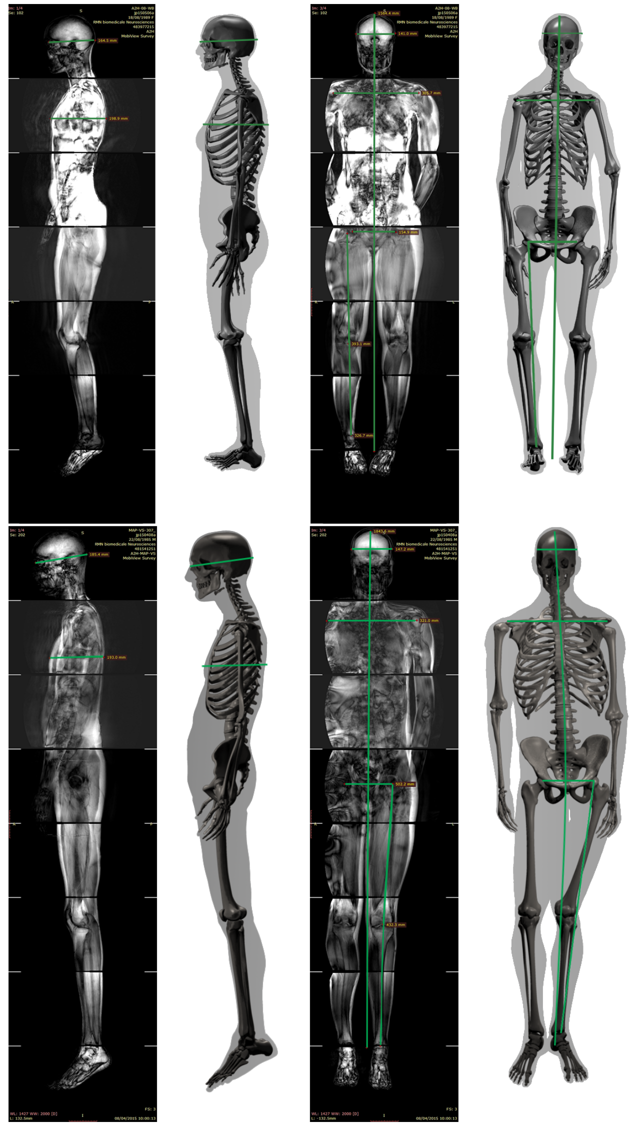

Figure 18: *Morphometric measurements (green lines) used to compare our results with ground truth MRI data. Top: 1.55m female. Bottom: 1.85m male.*

In *Kinect* body tracking skeleton data, we observe a lot of change in limbs lengths. Thus, we pinpoint anatomical specific points (long bones protuberences) onto the MRI and onto the user-specific associated 3D model. With these specific points, we compare ulna and tibia lengths in real and 3D data. We suffer an average 5.2% error in limb lengths, most of the time the user-specific 3D model lacks a few centimeters. This percentage seems quite acceptable taking into account *Kinect* raw data noisiness.

To evaluate the torso body part realism, we propose to compare the user-specific 3D model and full-body MRI data by comparing the distance between left and right humerus bone heads. The average error between 3D and MRI data is rather small: 1.5%.

Fig.*10* shows that the point cloud and skin are fairly close; our generic skin being registered for a woman or a man, but what about internal anatomy? We know that women hips are in average larger than men to allow birth. The 3.1% error between MRI and 3D data (the 3D data being always bigger than the MRI data) demonstrates that the distances between left and right femoral bones head difference between women and men is well transcribed in internal anatomy.

Using the lateral view, we pinpoint specific points to find rib cage depth. The user-specific rib cage is always bigger than the MRI data (around 20% bigger). It may be due to the difference in posture during acquisition (for MRI data, the user is lying whereas for *Kinect* data the user is standing). It may also be due to the use of a partial point cloud instead of a complete one. Due to front view capture, we observe depth errors in the skull as well: the skull is about 12% bigger in depth in 3D than in MRI data.

4.2 Validation with User Study

In a second set of experiments, are involved 20 different subjects with no motor problems and working everyday on tools involving medicine or medical imaging. For each subject, we captured a range of full body motions involving upper and lower limb motion as well as torso motion.
The group is composed of 13 men between 24 and 54 years old (average height: $181cm$, average weight: $82.6kg$), and 7 women between 22 and 44 years old (average height: $164cm$, average weight: $61.7kg$). Table 1 gives global informations about each subject.

Gender	Age	Body Mass Index	*Anatomy Knowledge*
male	44	32.4	*poor knowledge*
male	47	29.3	*professional*
male	31	27.4	*professional*
male	41	26.3	*poor knowledge*
female	44	26	*average knowledge*
male	41	25.5	*average knowledge*
male	24	25.1	*poor knowledge*
female	28	24.4	*professional*
male	30	24.3	*average knowledge*
male	30	24.3	*professional*
male	27	23.7	*average knowledge*
male	54	23.6	*average knowledge*
female	22	23.5	*poor knowledge*
female	34	22.2	*average knowledge*
female	32	22	*professional*
female	24	21.8	*poor knowledge*
male	29	21.6	*average knowledge*
male	28	21.6	*average knowledge*
female	31	21.3	*poor knowledge*
male	31	20.5	*average knowledge*

Table 1: *Global informations about each user study subject.*

This user study was designed to evaluate the believability of our system. We defined eight criteria which evaluation is given in table 2 to evaluate the quality of our mirror-like AR system.

Body position range (criterion *C01*) corresponds to motions while standing, crouching or sitting. In most cases, the results are well received. For other cases, limitations are directly connected to *Kinect* occlusion limitations.

Body orientation range (criterion *C02*) corresponds to body orientation from *Kinect* point of view: eg. facing, profile, 3/4, back. When *Kinect* raw data are occluded or self-occluded, our system returns false motion poses: the more occlusion in *Kinect* raw data, the more errors we will have. A major topic of future work is to be able to handle important occlusion zones.

Motion range (criterion *C03*) defines simple motions like Flexion/extension of the knee, as well as complex motions in the extremities like finger motion or supination/pronation of the arm. We obtain high motion quality for simple motions, for complex motions we are limited by *Kinect*: this criterion is still in need of improvements. The *Kinect* SDK outputs a small number of joints which limits the body motion possibilities (eg. spine bending). Head Tracking could be improved by using *Kinect* facial tracking.

For **Motion fluidity and delay** (criterion *C04*) and **Motion consistency** (criterion *C05*), the goal is reached. Motion consistency refers to the absence of outliers during motion. We should state the fact that part of the visual latency that might occur comes from the low frame rate of the color map display.

Motion plausibility (criterion *C06*) corresponds essentially to joint DOFs and angular limits. For this criterion we obtain different results depending on the body segment studied. For instance, it is more easy to implement constraint for 1DOF joints than for 3DOFs joints such as spine or shoulders joints due to motion range.

Anatomy realism (criterion *C07*) gives a feedback on the registration method by focusing on limb length and torso width. For this criterion, people with professional knowledge in anatomy were the only ones to rate the user-specific anatomy as average.

For almost everyone, the **Augmented reality** (criterion *C08*) results were of good level. The overall quality can even be increased with mesh texturing.

	C01	*C02*	*C03*	*C04*	*C05*	*C06*	*C07*	*C08*
$--$	0	3	1	0	0	2	0	0
$+-$	4	10	6	2	2	5	5	1
$++$	16	7	13	18	18	13	15	19

Table 2: *User study complied results according to the quality criteria for a mirror-like augmented reality system. For each criterion: number of user having a bad/average/good evaluation of the criterion.*

5 Conclusion

We present the first live system of personalized anatomy in motion. Superimposing the anatomy onto the user's image allows us to create a real-time augmented reality experience. The attached video (see *https://youtu.be/Ip17-Vaqqos*) illustrates the application pipeline and shows AR results of our system. We believe that the basic *Kinect* body tracking enhanced with our method is sufficiently accurate for our needs.

The system could be extended in different ways. Most users claimed in the user study that the overall quality of AR is of good level. However, some artefacts are still visible during motion and future work will be done to ensure that the 3D user-specific data always lie within the user's silhouette; we could apply a silhouette retargeting as in [Zhou et al. 2010] to correct our hierarchical body tracking system.
The addition of biomechanical simulations could allow to get more realistic deformations of soft tissues and organs but this could be at the cost of interactivity.
To show full body muscular activity for every possible body motion, inverse dynamics [Murai et al. 2010] will also be developed.
An improvement in the skin registration can be done by reducing

the 9DOFs controllers to 6DOFs (3 rotations and 3 scales). This can be done by exploiting appropriately the hierarchical structure and would allow more robustness and skin consistency around body joints.

Our work is designed to be used as a tool for anatomy learning for medical and sports students. This is why in the future it is planned to display anatomical educational content (text, images, videos, *etc*) in addition to the AR visualization.
This system could also be used as a way to communicate between medical practionners and their patients, about surgery, rehabilitation or any other health issue. Novel artistic content might also be produced using our technology, as well as interactive advertising.

Our system has been featured as a live demo during two conferences and at the *Consumer Electronic Show*. More than 400 people have been able to test it out. Most of them enjoyed the experience and a lot of them recommended it and came back with others.

Acknowledgments

Many thanks to Estelle Charleroy our Graphist Artist, and to Olivier Carre for engineering advices. We also would like to thanks all people involved in the user study, and all people helped testing the application.
MRI data used for validation are from the A2H study promoted by Grenoble University Hospital.
This work has been partially supported by the LabEx PERSYVAL-LAB (ANR-11-LABX-0025).

References

ASTERIADIS, S., CHATZITOFIS, A., ZARPALAS, D., ALEXIADIS, D. S., AND DARAS, P. 2013. Estimating human motion from multiple kinect sensors. In *Proceedings of the 6th International Conference on Computer Vision/Computer Graphics Collaboration Techniques and Applications*, ACM, 3.

BĂLAN, A. O., AND BLACK, M. J. 2008. The naked truth: Estimating body shape under clothing. In *Computer Vision–ECCV 2008*. Springer, 15–29.

BAUER, A., PACLET, F., CAHOUET, V., DICKO, A.-H., PALOMBI, O., FAURE, F., AND TROCCAZ, J. 2014. Interactive visualization of muscle activity during limb movements: Towards enhanced anatomy learning. In *Eurographics Workshop on Visual Computing for Biology and Medicine,(VCBM) 2014*.

BAUER, A., DICKO, A.-H., PALOMBI, O., FAURE, F., AND TROCCAZ, J. 2015. Living book of anatomy project : See your insides in motion! In *SA'15 Emerging Technologies, November 02-06, 2015, Kobe, Japan*.

BICHLMEIER, C. 2010. *Immersive, Interactive and Contextual In-Situ Visualization for Medical Applications*. PhD thesis, Technische Universitt Mnchen (TUM), Chair for Computer Aided Medical Procedures.

BLUM, T., KLEEBERGER, V., BICHLMEIER, C., AND NAVAB, N. 2012. mirracle: An augmented reality magic mirror system for anatomy education. In *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*, IEEE, 115–116.

BORNER, A., AND KIRSCH, S., 2015. Anatomie spiegel. http://www.qwertzus.com/anatomie-spiegel/.

DAO, N.-L., DENG, T., AND CAI, J. 2014. Fast and automatic body circular measurement based on a single kinect. In *Asia-Pacific Signal and Information Processing Association, 2014 Annual Summit and Conference (APSIPA)*, IEEE, 1–4.

DICKO, A.-H., TIANTIAN, L., GILLES, B., KAVAN, L., FAURE, F., PALOMBI, O., AND CANI, M.-P. 2013. Anatomy transfer. *ACM Transactions on Graphics (TOG) 32*, 6, 188.

EBERLY, D., 2008. Constrained quaternions using euler angles, July.

ESPITIA-CONTRERAS, A., SANCHEZ-CAIMAN, P., AND URIBE-QUEVEDO, A. 2014. Development of a kinect-based anthropometric measurement application. In *Virtual Reality (VR), 2014 iEEE*, IEEE, 71–72.

FAURE, F., GILLES, B., BOUSQUET, G., AND PAI, D. K. 2011. Sparse Meshless Models of Complex Deformable Solids. *ACM Transactions on Graphics 30*, 4 (July), Article No. 73.

FAURE, F., DURIEZ, C., DELINGETTE, H., ALLARD, J., GILLES, B., MARCHESSEAU, S., TALBOT, H., COURTECUISSE, H., BOUSQUET, G., PETERLIK, I., AND COTIN, S. 2012. SOFA: A Multi-Model Framework for Interactive Physical Simulation. In *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery*, Y. Payan, Ed. Springer. http://www.sofa-framework.org.

GILLES, B., REVERET, L., AND PAI, D. 2010. Creating and animating subject-specific anatomical models. *Computer Graphics Forum* (June), http://onlinelibrary.wiley.com/doi/10.1111/j.1467–8659.2010.01718.x/abstract.

GILLES, B., BOUSQUET, G., FAURE, F., AND PAI, D. 2011. Frame-based elastic models. *ACM Transactions on Graphics 30*, 2 (Apr.), Article No. 15.

GILLES, B., FAURE, F., BOUSQUET, G., AND PAI, D. K. 2013. Frame-based interactive simulation of complex deformable objects. In *Deformation Models*, M. Hidalgo González, A. Mir Torres, and J. Varona Gómez, Eds., vol. 7 of *Lecture Notes in Computational Vision and Biomechanics*. Springer, 145–166.

GONZÁLEZ, A., HAYASHIBE, M., AND FRAISSE, P. 2012. Estimation of the center of mass with kinect and wii balance board. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, IEEE, 1023–1028.

GUAN, P., FREIFELD, O., AND BLACK, M. J. 2010. A 2d human body model dressed in eigen clothing. In *Computer Vision–ECCV 2010*. Springer, 285–298.

HELTEN, T., BAAK, A., BHARAJ, G., MULLER, M., SEIDEL, H.-P., AND THEOBALT, C. 2013. Personalization and evaluation of a real-time depth-based full body tracker. In *3D Vision-3DV 2013, 2013 International Conference on*, IEEE, 279–286.

HSIAO, S.-W., AND CHEN, R.-Q. Applying kinect on the development of a customized 3d mannequin. *Evaluation 39*, 61530.

KAMPHUIS, C., BARSOM, E., SCHIJVEN, M., AND CHRISTOPH, N. 2014. Augmented reality in medical education? *Perspectives on medical education 3*, 4, 300–311.

LI, H., VOUGA, E., GUDYM, A., LUO, L., BARRON, J. T., AND GUSEV, G. 2013. 3d self-portraits. *ACM Transactions on Graphics (Proceedings SIGGRAPH Asia 2013) 32*, 6 (November).

LUSK, M. M., AND ATKINSON, R. K. 2007. Animated pedagogical agents: Does their degree of embodiment impact learning from static or animated worked examples? *Applied cognitive psychology 21*, 6, 747–764.

MACEDO, M. C., APOLINÁRIO JR, A. L., SOUZA, A. C., AND GIRALDI, G. A. High-quality on-patient medical data visualization in a markerless augmented reality environment.

MAITRE, X., 2014. Newscientist : Digital mirror reveals what lies under your skin, April.

MALINOWSKI, M. J., AND MATSINOS, E. 2015. On using the microsoft kinect sensors to determine the lengths of the arm and leg bones of a human subject in motion. *eprint arXiv:1505.00371 [physics.med-ph]*.

MENG, M., FALLAVOLLITA, P., BLUM, T., ECK, U., SANDOR, C., WEIDERT, S., WASCHKE, J., AND NAVAB, N. 2013. Kinect for interactive ar anatomy learning. In *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on*, IEEE, 277–278.

MERCIER-GANADY, J., LOTTE, F., LOUP-ESCANDE, E., MARCHAL, M., AND LÉCUYER, A. 2014. The mind-mirror: See your brain in action in your head using eeg and augmented reality. In *Virtual Reality (VR), 2014 iEEE*, IEEE, 33–38.

MURAI, A., KUROSAKI, K., YAMANE, K., AND NAKAMURA, Y. 2010. Musculoskeletal-see-through mirror: Computational modeling and algorithm for whole-body muscle activity visualization in real time. *Progress in Biophysics and Molecular Biology 103*, 2-3, 310–317.

NORDIN, M., AND FRANKEL, V. H. V. H. 2001. *Basic biomechanics of the musculoskeletal system*. Lippincott Williams & Wilkins, Philadelphia (Pa.).

PFISTER, A., WEST, A. M., BRONNER, S., AND NOAH, J. A. 2014. Comparative abilities of microsoft kinect and vicon 3d motion capture for gait analysis. *Journal of medical engineering & technology 38*, 5, 274–280.

PIOTIN, S., BENASSAROU, A., BLANCHARD, F., NOCENT, O., AND BERTIN, E. 2013. Abdominal morphometric data acquisition using depth sensors. In *e-Health Networking, Applications & Services (Healthcom), 2013 IEEE 15th International Conference on*, IEEE, 653–657.

QUAH, C. K., GAGALOWICZ, A., ROUSSEL, R., AND SEAH, H. S. 2005. 3d modeling of humans with skeletons from uncalibrated wide baseline views. In *Computer Analysis of Images and Patterns*, Springer, 379–389.

SAITO, S., ZHOU, Z.-Y., AND KAVAN, L. 2015. Computational bodybuilding: Anatomically-based modeling of human bodies. *ACM Trans. Graph. 34*, 4.

SHEN, W., DENG, K., BAI, X., LEYVAND, T., GUO, B., AND TU, Z. 2012. Exemplar-based human action pose correction and tagging. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, IEEE, 1784–1791.

SOTIRAS, A., DAVATZIKOS, C., AND PARAGIOS, N. 2012. Deformable Medical Image Registration: A Survey. Research Report RR-7919, INRIA, Sept.

VLASIC, D., BARAN, I., MATUSIK, W., AND POPOVIĆ, J. 2008. Articulated mesh animation from multi-view silhouettes. In *ACM Transactions on Graphics (TOG)*, vol. 27, ACM, 97.

WEI, X., ZHANG, P., AND CHAI, J. 2012. Accurate realtime full-body motion capture using a single depth camera. *ACM Transactions on Graphics (TOG) 31*, 6, 188.

ZENG, M., CAO, L., DONG, H., LIN, K., WANG, M., AND TONG, J. 2015. Estimation of human body shape and cloth field in front of a kinect. *Neurocomputing 151*, 626–631.

ZHOU, S., FU, H., LIU, L., COHEN-OR, D., AND HAN, X. 2010. Parametric reshaping of human bodies in images. In *ACM SIGGRAPH 2010 Papers*, ACM, New York, NY, USA, SIGGRAPH '10, 126:1–126:10.

ZHOU, L., LIU, Z., LEUNG, H., AND SHUM, H. P. 2014. Posture reconstruction using kinect with a probabilistic model. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, ACM, 117–125.

ZHU, L., HU, X., AND KAVAN, L. 2015. Adaptable anatomical models for realistic bone motion reconstruction. *Comput. Graph. Forum 34*, 2.